Cellular and metabolic characteristics of pre-leukemic hematopoietic progenitors with GATA2 haploinsufficiency

Avigail Rein,[1,2,3] Ifat Geron,[1,2,3,4] Eitan Kugler,[1,2,3] Hila Fishman,[1,2,3] Eyal Gottlieb,[5] Ifat Abramovich,[5] Amir Giladi,[6] Ido Amit,[6] Roger Mulet-Lazaro,[7] Ruud Delwel,[7,8] Stefan Gröschel,[7,9,10] Smadar Levin-Zaidman,[11] Nili Dezorella,[11] Vered Holdengreber,[12] Tata Nageswara Rao,[13] Joanne Yacobovich,[2] Orna Steinberg-Shemer,[2,4] Qiu-Hua Huang,[14] Yun Tan,[14,15] Sai-Juan Chen,[14,15] Shai Izraeli[1,2,3,4] and Yehudit Birger[1,2,3,4]

[1]Department of Human Molecular Genetics and Biochemistry, Sackler Medical School, Tel Aviv University, Tel Aviv, Israel; [2]The Rina Zaizov Division of Pediatric Hematology-Oncology, Schneider Children's Medical Center, Petah Tikva, Israel; [3]Functional Genomics and Childhood Leukaemia Research, Sheba Medical Centre, Tel-Hashomer, Israel; [4]Felsenstein Medical Research Center, Sackler School of Medicine, Tel Aviv University, Petah Tikva, Israel; [5]Technion Integrated Cancer Center, Faculty of Medicine, Technion Israel Institute of Technology, Haifa, Israel; [6]Department of Immunology, Weizmann Institute of Science, Rehovot, Israel; [7]Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands; [8]Oncode Institute, Erasmus University Medical Center, Rotterdam, The Netherlands; [9]Molecular Leukemogenesis, Deutsches Krebsforschungszentrum, Heidelberg, Germany; [10]Department of Internal Medicine V, Heidelberg University Hospital, Heidelberg, Germany; [11]Electron Microscopy Unit, Weizmann Institute of Science, Rehovot, Israel; [12]Electron Microscopy Unit, IDRFU, Faculty of Life Sciences, Tel Aviv University, Israel; [13]Stem Cells and Leukemia Laboratory, University Clinic of Hematology and Central Hematology, Department of Biomedical Research (DBMR), Inselspital Bern, University of Bern, Bern, Switzerland; [14]State Key Laboratory of Medical Genomics, Shanghai Institute of Hematology, Rui Jin Hospital Shanghai, China and [15]Jiao Tong University School of Medicine, Shanghai, China

Correspondence:
S. Izraeli
sizraeli@gmail.com

Y. Birger
Yehudit.Birger@gmail.com

Received: June 21, 2021.
Accepted: December 1, 2022.
Early view: December 7, 2022.

https://doi.org/10.3324/haematol.2022.279437

Supplemental Data

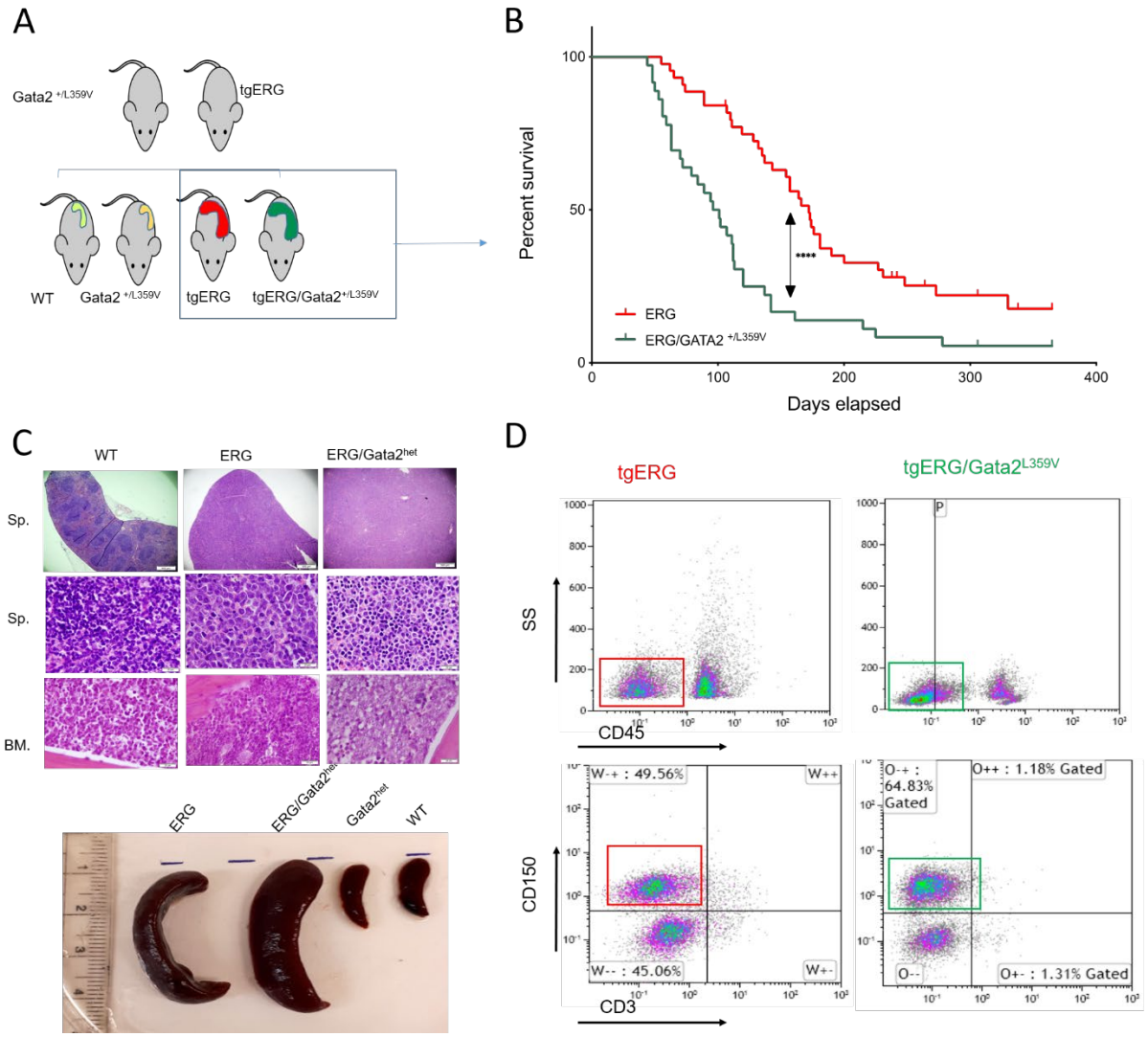
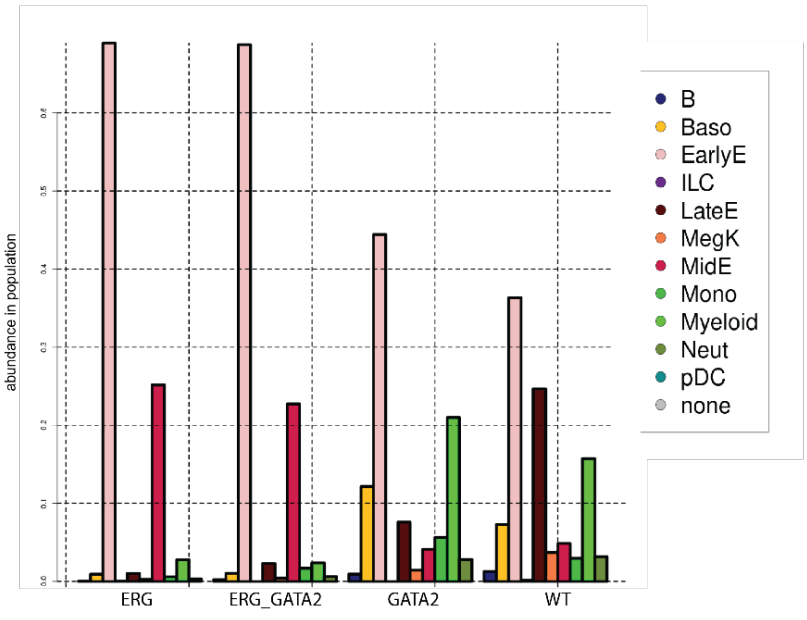
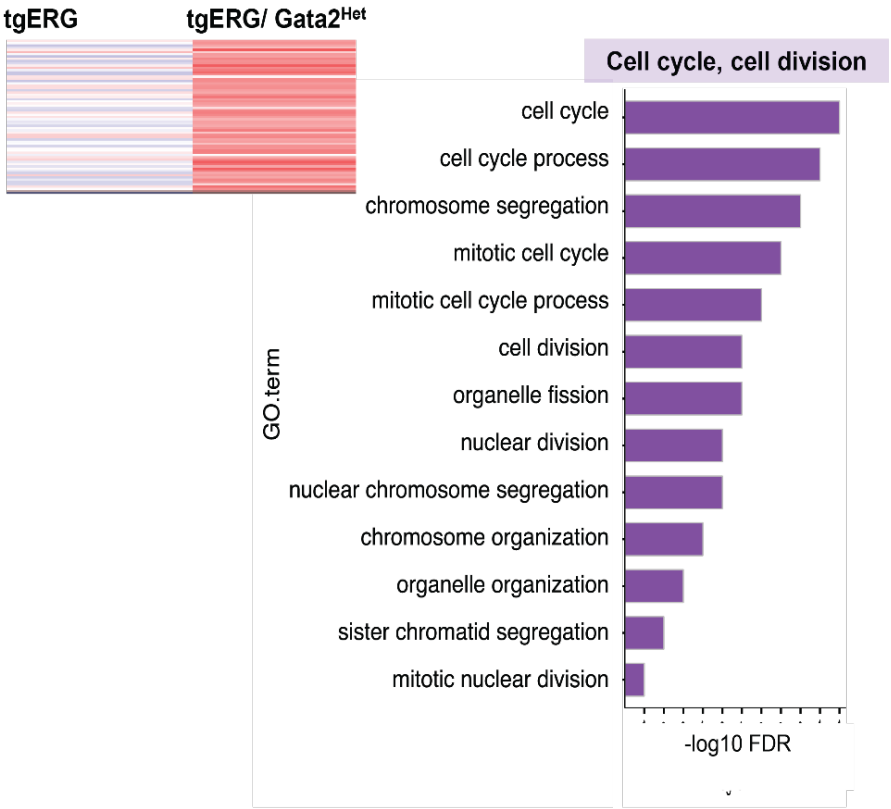


Figure. S1. **Accelerated leukemia and reduced survival in ERG/Gata2$^{+/L359V}$ mice. A.** Gata2$^{+/L359V}$ and tgERG mice mating strategy generates 4 different genotypes. **B.** tgERG/Gata2$^{+/L359V}$ double mutated mice, had shortened survival compared with tgERG littermates ($p<0.0001$, log rank Mantel-Cox test). **C.** Histopathological sections of tgERG/Gata2^{het} and tgERG leukemic mice. (up)-H & E stained histopathological sections of spleen and bone marrow demonstrated infiltrations of leukemic cells, (light microscope: upper panel 4X, middle and bottom panels 100X magnifications, scale bars calibrated as shown). (bottom) - Gross appearance of leukemia infiltrated spleens, alongside to normal sized spleens. Sp.- spleen BM- Bone marrow. **D.** Immunophenotype profile of the leukemia consisted of lineage negative, CD45^{dim} cells, expressing high level of CD150

A



B



C

Aspm
Lbr
Cenpf
Nsl1
Mastl
Ckap5
Kif18a
Casc5
Wdr76
Cep152
Ncaph
Tpx2
Pola1
Kif4
Trim59
Golim4
Gabpb2
Hist2h2ac
Cenpe
Depdc1a
Melk
Smc2
Cdk5rap2
Mast2
Mtf1
Clspn
Dbf4
Ncapg
Pi4k2b
Cenpc1
Helq
Cit
Kntc1
Srrt
Rad18
Fancd2
Ccdc77
Ncapd2
Rad51ap1
Zfp626
Fanci
Prc1
Blm
Relt
Nup98
Rrm1
Ccp110
Ikzf5
Mki67
Mtfr2
Parpbp

Esco2
Anln
Ncapd3
Chek1
Kif23
Leo1
Ttk
Cep70
Cnot10
Kif15
Hmmr
Zfp672
Zfp39
Mis12
Spag5
Atad5
Cltc
Prr11
Top2a
Brca1
Idi1
Hist1h1b
Hist1h3i
Hist1h2ap
Hist1h2bj
Hist1h2ae
Hist1h4d
Hist1h1e
Hist1h3c
Hist1h3a
Hist1h1a
Hist1h2al
Zfp367
Cenph
Gen1
G2e3
Fancm
Mis18bp1
Ccdc88c
Smek1
Ncapg2
Rictor
6030458C11R
ik
Rad21
Senp1
Cbx5
Senp2
Pcyt1a
C330027C09R
ik
Pdpk1
Kat2b
Sgol1
Taf7
Cep76
Smc5
Fam122a
Kif11

Figure. S2. **10X RNAseq analysis reveals skewed differentiation and enhanced proliferation signature in tgERG/Gata2^{het} HSPCs**. **A.** Bar graph depicting abundance of population differentiation as reflected by gene expression pattern in the tested genotypes. **B.** illustration of k-means clustering of 10X RNAseq data, showing tgERG/Gata2^{het} HSPC highly express gene associated with 'Cell Cycle', and 'Cell Division' GO terms (K-means, an unsupervised algorithm that is used for cluster identification in a data set, was calculated with STRING platform[1]). **C.** List of genes clustered by k-means.

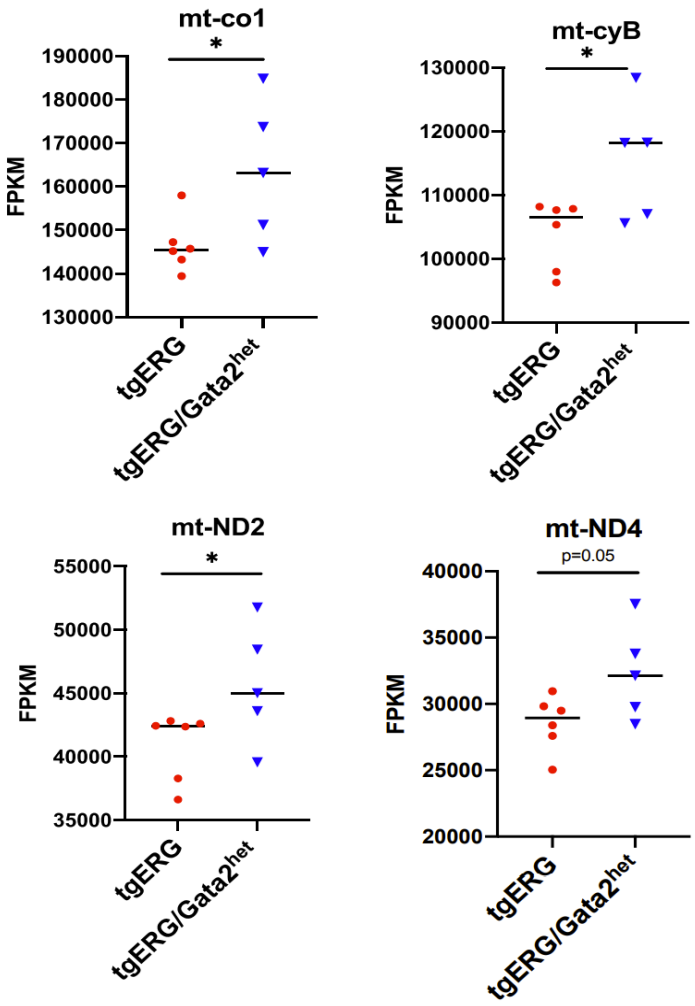


Figure S3. Selected mitochondrial genes upregulated in RNAseq of tgERG/Gata2^{het} pre-leukemia HSPCs. Individual mitochondrial genes (mt-co1: (Mitochondrially Encoded Cytochrome C Oxidase I),mt-cyB (Mitochondrially Encoded Cytochrome B), mt-ND2 (Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 2), mt-ND4 (Mitochondrially Encoded NADH:Ubiquinone Oxidoreductase Core Subunit 4) were differentially expressed in tgERG/Gata2^{het} samples (blue, n=5) compared with tgERG (red, n=6). (student t-test $p<0.05$, and in mtND4=0.05) FPKM- Fragments Per Kilobase of transcript per Million mapped reads.

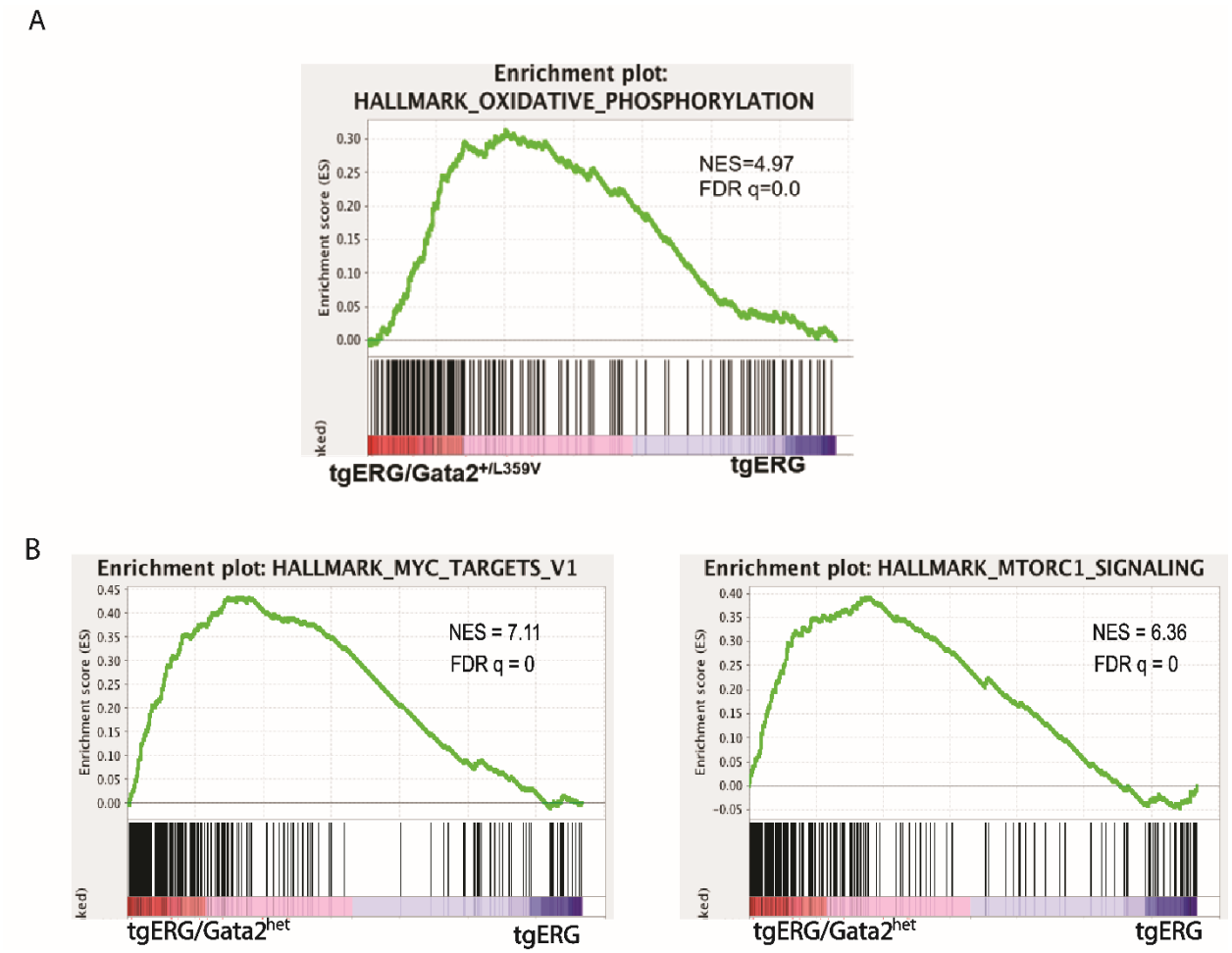


Figure S4. Expression signature enrichment in transgenic mice leukemic cells: A Gene set enrichment analysis of tgERG/Gata2+/L359V leukemic cells reviles that ‘Oxidative Phosphorylation’ pathway genes are enriched by gERG/Ga a2$^{+/L359V}$ expression. B Gene set enrichment analysis of ERG/Gata2^{het} leukemic cells revieles enrichment of MYC and mTOR signaling pathways genes (NES= normalized enrichment score, FDR= false discovery rate).

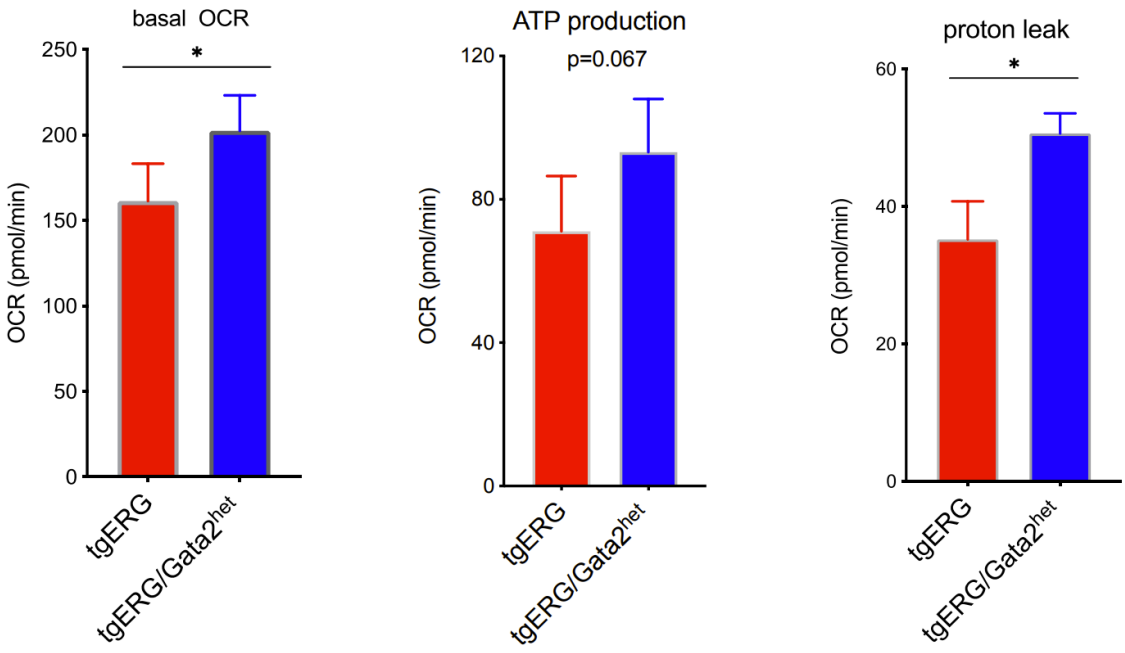


Figure S5. Increased Oxygen Consumption Rates of tgERG/Gata2^{het} HSPCs.

Basal oxygen consumption rates are significantly higher in tgERG/Gata2^{het} pre-leukemic HSPCs (**left**), and a trend (**middle**) of higher ATP production is also noted. Proton leak is shown to be higher in tgERG/Gata2^{het}, suggestive of an uncoupling effect of the respiratory chain (**right**). ($p<0.05$ n=3, two-tailed paired student t-test).

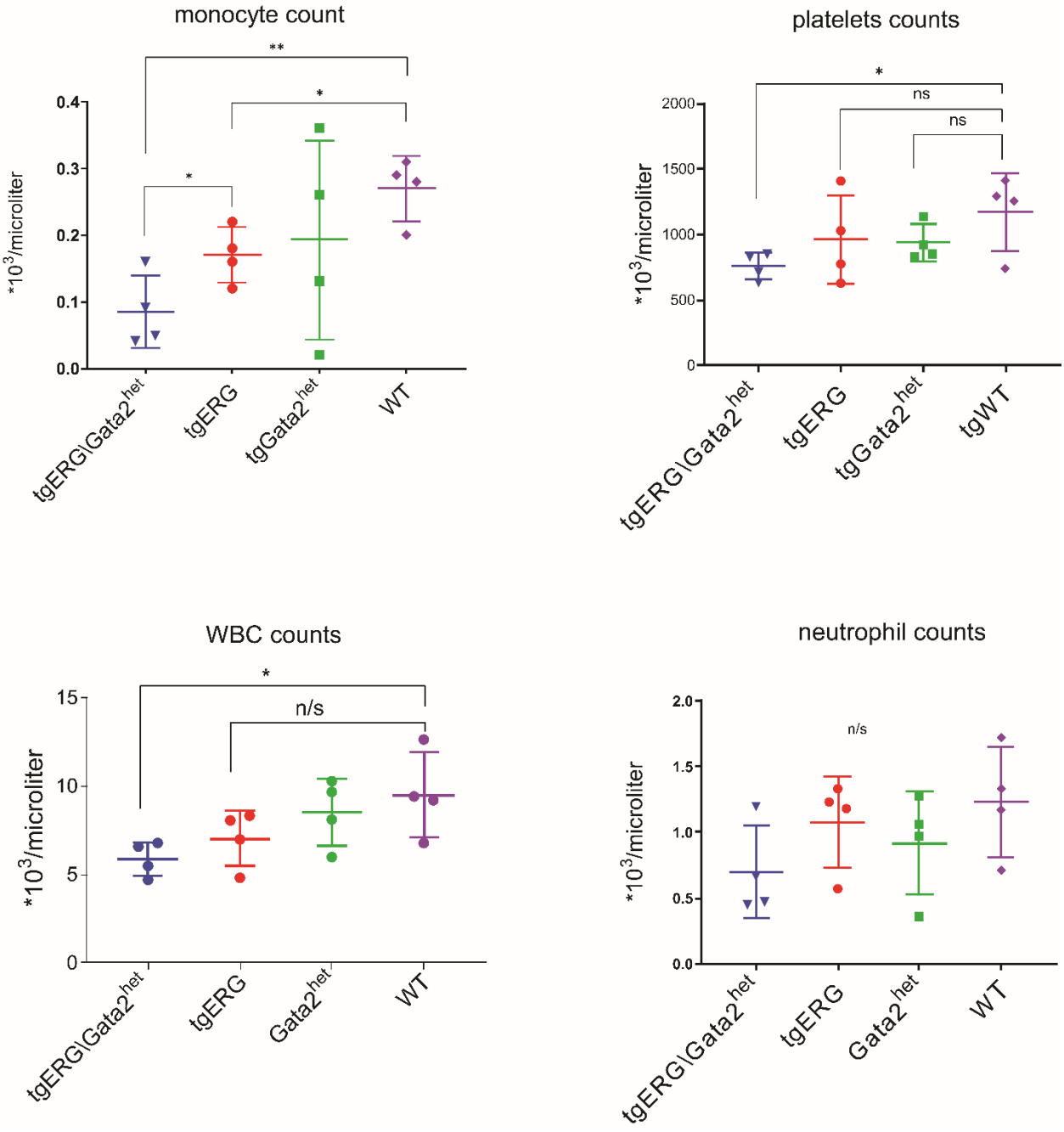


Figure. S6. tgERG/Gata2^{het} preleukemic mice develop blood cytopenia before clinical signs appear. Peripheral blood counts of 16 (4 of each genotype) preleukemic age- matched 6weeks old mice, showing that tgERG/Gata2^{het} encounter a decrease in several blood lineages before clinical signs are apparent. ($p<0.05$; one-way ANOVA).

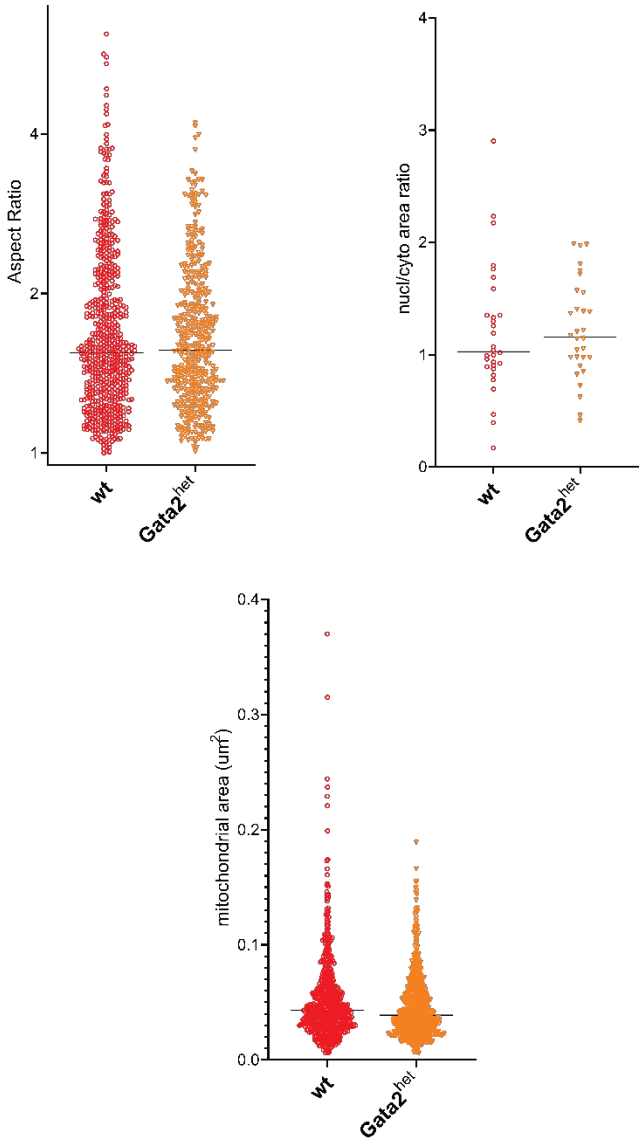


Figure S7. Lin- GATA2^{het} have normal mitochondrial morphology.

Transmission electrons microscopy captures imaging Analyses (NIH ImajeJ Fiji tool (50)) of lin- cells from BM of Gata2^{het} and wt 5 weeks old littermates ($n_{(mice)}$=3 $n_{(cells/mouse\ sample)}$=10 for both Gata2het and wt) Dot plots depicting aspect ratio (left) nucleos to cytoplasm area ratio (right) and mitochondrial area (bottom). No significant difference could be found in the measured mitochondrial parameters. (unpaired student T- test).

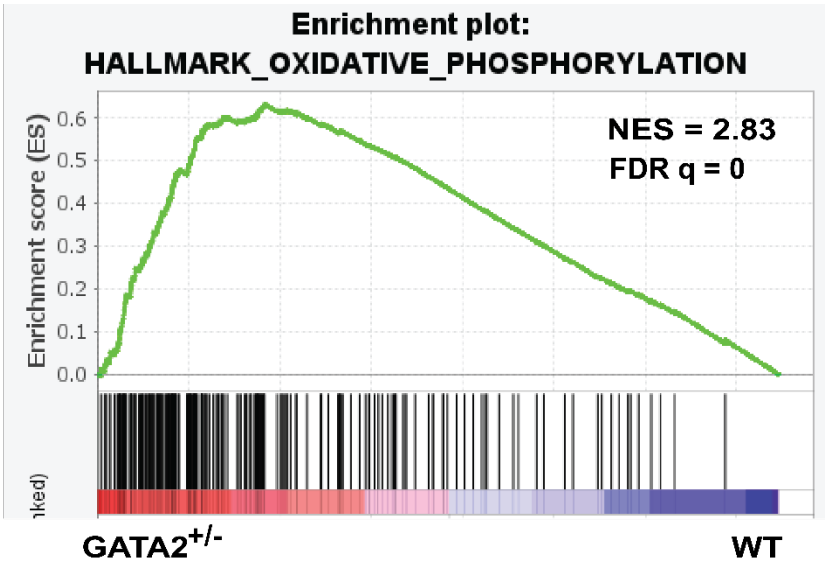


Figure S8. Oxidative Phosphorylation expression signature enrichment in GATA2het cells: Erythroid progenitor cells from GATA2het and wt background were subjected to RNA sequencing. Gene set enrichment analysis reveals enrichment in oxidative phosphorylation pathway genes.

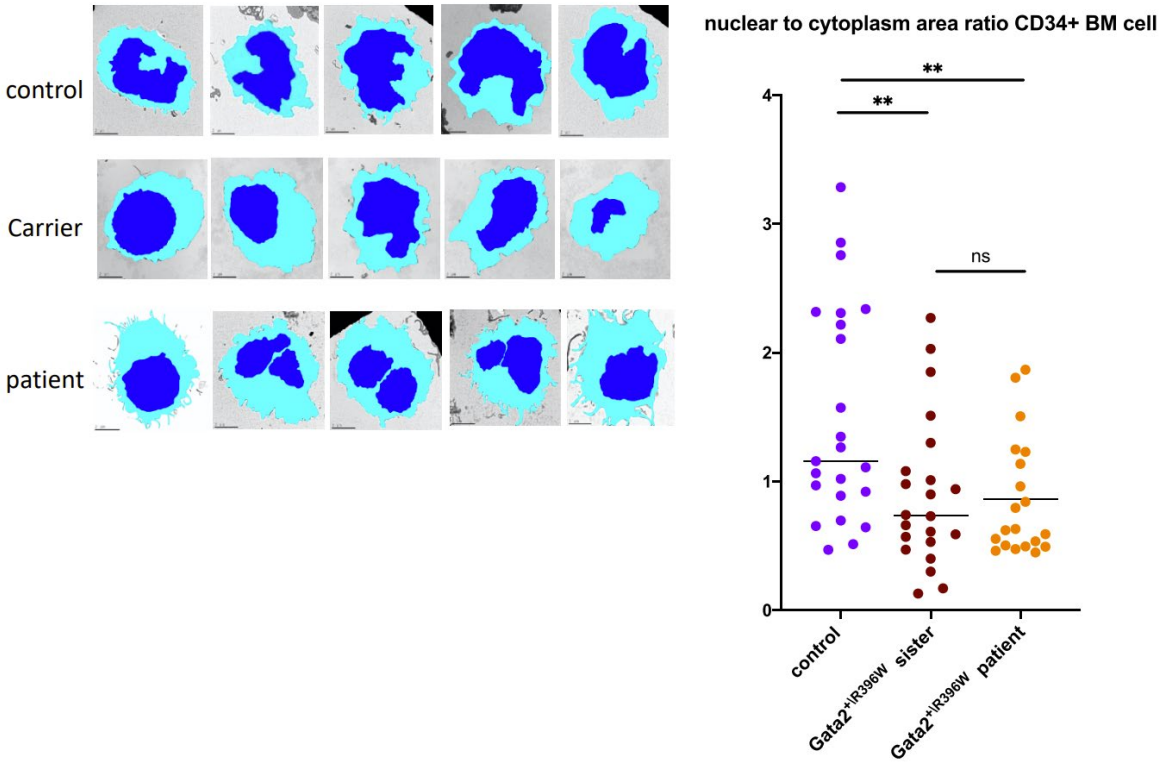


Figure S9. GATA2$^{+/R396W}$ CD34+ HSPCs display decreased nuclear to cytoplasm ratio. TEM imaging of GATA2$^{+/R396W}$ CD34+ HSPCs were analyzed using ImajeJ Fiji tool. Left - representative view of Nuclear to Cytoplasmic Ratio. Demarcation of cytoplasm (light blue) and nuclei (dark blue). CD34+ HSPCs from the GATA2$^{+/R396W}$ patient (lower panel), carrier sister (middle) and healthy donor (upper panel) were analyzed. (1:4 800-6 800 magnification scale, calibrated by Fiji tool). Right- Dot plots depicting nuclear to cytoplasm area ratio of CD34+ BM cells. Lines represent mean. A decreased nuclear to cytoplasmic ratio (NCR) was found in both the patient (orange) and the carrier sister (brown) compared to healthy donor (purple) (multiple student t-tests, n= $p<0.0001$).

Table S1- Flow cytometry antibodies

Target protein	Company	Catalog number	Conjugate	reactivity
cKit (CD117)	eBioscience	25-1172-82	PE-Cy5	Mouse
Sca1	eBioscience	56-5981-82	Alexa fluor 700	Mouse
Gr-1	Biolegend	108412	APC	Mouse
CD11b	BD	553311	PE	Mouse
CD41	Biolegend	133905	PE	Mouse
CD61	Biolegend	104315	APC	Mouse
CD150	Biolegend	115909	APC	Mouse
CD4	eBioscience	17-0041-82	APC	Mouse
SCA1	eBioscience	56-5981-82	Alexa fluor 700	Mouse
GR-1	Biolegend	108412	APC	Mouse
CD3	BD	555275	PE	Mouse
CD45	Biolegend	103114	PE-CY7	Mouse
CD8	Biolegend	100714	APC-CY7	Mouse
CD34	eBioscience	48-0341-82	efluor450	Mouse
7AAD	BD	51-68981		Mouse/Human
CD48	Biolegend	103418	Pacific blue	mouse

Table S2

Forward and reverse primers used for mtDNA/nDNA ratio:

16S rRNA forward and reverse primers

FWD: 5 -CCGCAAGGGAAAGATGAAAGAC-3

REV: 5 -TCGTTTGGTTTCGGGGTTTC-3

ND1 forward and reverse primers

FWD: 5 -CTAGCAGAAACAAACCGGGC-3

REV: 5 -CCGGCTGCGTATTCTACGTT-3

HK2 forward and reverse primers

FWD: 5 -GCCAGCCTCTCCTGATTTTAGTGT-3

REV: 5 -GGGAACACAAAAGACCTCTTCTGG-3

Supplemental Methods

Mouse models

ERG Tg mice model:

The *ERG* transgenic mice were generated by subcloning a 1.4-kb fragment of human ERG3 hematopoietic isoform into the HS21/45-*vav* vector [1]. This construct was microinjected into pronuclei of fertilized C57BL6 F1 oocytes and several founders were identified by PCR. All animals expressing ERG developed acute leukemia and died by the age of 7 months. Approximately 70% of leukemia were characterized as AML and approximately 30% as T-ALL [2]. DNA from tail clip of born mice is analyzed at the age of 4 weeks for the presence of human ERG by PCR.

$Gata2^{L359V}$ mice model:

Gata2-L359V knock-in murine model was generated by introducing the L359V mutation into the ZF2 domain, which is located in the exon 5 of the murine Gata2 gene. The Gata2-L359V knock-in mice were generated in C57BL/6 mice and on BALB/c background through backcross breeding. Gata2-L359V homozygous mutation resulted in embryonic lethality around E11.5 due to defects in primitive erythropoiesis and severe anemia. Gata2-L359V heterozygous adult mice exhibited defective hematopoietic development and block in the differentiation of chronic myeloid leukemia (CML) cells [3]. DNA from tail clip of born mice is analyzed at the age of 4 weeks for the presence of Gata2-L359V mutation by PCR.

$GATA2^{het}$ mice model:

Neomycin-resistant cassette replaced the carboxyl zinc finger domain of the murine GATA2 gene in a targeting vector with GATA2 complementary DNA sequence, generating a null mutation. Homozygous mice for GATA2 null died around E11.5. Mice

heterozygous for GATA2 appeared normal [4]. DNA from tail clip of born mice is analyzed by PCR at the age of 4 weeks for the presence of Neo sequence that is present in the knock-in construct.

tgERG/Gata2^{L359V} mice model:

Double transgenic mice were generated by crossing heterozygous TgERG mice [2] with *Gata2$^{+/L359V}$ knock-in* mice [3]. Four genotypes were identified by PCR analysis of DNA from tail clip of born mice at the age of 4 weeks using specific primers for Gata2-L359V mutation and specific primers for human ERG sequence - WT, tgERG, Gata2 L359V and the double transgene tgERG/Gata2 L359V.

tgERG/GATA2^{het} mice model:

Double transgenic mice were generated by crossing heterozygous TgERG mice [2] with *Gata2*$_{het}$ mice [4] . Four genotypes were identified by PCR analysis of DNA from tail clip of born mice at the age of 4 weeks using specific primers for the Neo sequence that is present in the knock-in construct and specific primers for human ERG sequence - WT, tgERG, Gata2 L359V and the double transgene tgERG/Gata2 L359V).

10X RNA sequencing and analysis

scRNA expression was obtained by using 10X to prepare cDNA library (Single Cell 3' Reagent Kits v2) the library was then sequenced on Illumina NextSeq 500 (paired end sequencing) Cell Ranger 3.0 (http://10xgenomics.com) was used to process Chromium single cell RNA-seq output, using mouse genome mm10 (version 1.2.0) as reference. The Metacell pipeline was used to derive informative genes and compute cell-to-cell similarity, K-nn graph covers and derive distribution of RNA in cohesive groups of cells and to derive strongly separated clusters using bootstrap analysis and computation of graph covers on resampled data. Default parameters were used unless otherwise stated. Cells with less than 500 UMI were discarded for low quality, retaining 12312 cells for further analysis, with a median 10010 UMI per cell. Raw data is available on GEO through accession GSE143308.

Oxygen consumption analysis

Oxygen Consumption Rates (OCR) was measured using the Seahorse XF96 Analyzer (Agilent Technologies, Santa Clara, California, USA). Cells were suspended in XF Assay Medium supplemented with 10 mM glucose,1 mM pyruvate, 2mM glutamine and 100 µM BSA-Oleate conjugate and seeded in a Seahorse XF96 cell culture plate (30-µl volume) precoated with Cell-Tak (Fisher Scientific). Cells were left to adhere for a minimum of 30 min in a CO2- free incubator at 37 °C, after which 150 µl of XF Assay Medium was added into each well. The plate was left to equilibrate for 20 min in the CO2-free incubator before being transferred to the Seahorse XF96 analyzer. Initial calibration experiments were performed to determine cell density, oligomycin and FCCP concentrations.

Optimal cell density was set to 200000 cells/well. Measurement of OCR was done at baseline and following sequential injections of (i) oligomycin (2 µM), an ATP synthase inhibitor, (ii) carbonyl cyanide-4 phenylhydrazone (FCCP) (20 µM), a mitochondrial uncoupler, and (iii) antimycin A (1 µM) and rotenone (1 µM), complex III and complex I inhibitors, respectively. All reagents were purchased from Sigma-Aldrich, St. Louis, Missouri, USA, aliquoted in DMSO in X1000 and diluted in assay media in the day of running the assay. OCRs were normalized by cell number.

RNAseq

For preparation of libraries the polyA fraction (mRNA) was purified from total RNA following by fragmentation and generation of double stranded cDNA. Then, end repair, A base addition, adapter ligation and PCR amplification steps were performed. Libraries were evaluated by Qubit (Thermo fisher scientific) and TapeStation (Agilent). Sequencing libraries were constructed with barcodes to allow multiplexing of 28 samples in 4 lanes. Around 22-28 million single-end 60bp reads were sequenced per sample on Illumina HiSeq 2500 V4 instrument. Poly-A/T

stretches and Illumina adapters were trimmed from the reads using cut adapt. Resulting reads shorter than 30bp were discarded.

Reads were mapped to Mus Musculus reference genome GRCm38 using STAR v2.4.2a, supplied with gene annotations downloaded from Ensemble release 82. Expression levels for each gene were quantified using HTseqcount. Differential expression analysis was performed using DESeq2. Raw P values were adjusted for multiple hypothesis using the Benjamini-Hochberg method. Raw data is available in GEO through accession GSE143238.

Immunophenotyping

BM cells were washed and stained with fluorochrome conjugated antibodies. Cells were washed and analyzed by Gallios 3 laser/10 colors Flow Cytometer (Beckman coulter, Brea, California, USA).

Leukemia panels: Progenitors: Lin(-) C-Kit, CD45, CD34 CD150, CD48, and Sca1; myeloid: CD13, CD11b CD14, GR-1 ; MegaErythroid: CD41 and CD61 together with Ter-119 ; T-Lymphoid: CD3, CD4, CD8, CD44 and CD25. Viability was assessed with 7AAD and dead cells were excluded by Kaluza software.

Cell proliferation assay

Tracing of cells generations was done using dye dilution assay using CFSE (Fisher Scientific, Waltham, Massachusetts, USA, #C34554), and analyzed using flow cytometry at 0h, 24h, 48h, 72h, and 96h. Dilution rate was calculated by Mean Fluorescence Intensity, and translated into Proliferation Index (P.I) using ModFit LT software (Verity Software House, Topsham, Maine, USA).

Gene Set Enrichment Analysis (GSEA)

PreRanked gene list analysis was performed. All the genes received per sample were scored by log10 P-values. Plus-sign was assigned for upregulated genes and

minus for downregulated genes. MSigDB: Molecular Signature DataBase of GSEA was used. The run was on classic mode. 500 permutations were defined.

Statistical analysis

All data are presented as mean and standard error of the mean (SEM). Statistical analyses were performed using GraphPad Prism® 8(La Jolla, CA, USA). Statistical significance of the data was assessed by two tailed Students *t*-test, or Mann-Whitney for two groups, and one way ANOVA, or Kruskal Wallis for more than two groups. Kaplan Meyer survival curves were generated by Log-Rank Mantel-Cox test. A *P*-value of <0.05 was considered significant.

Transmission electron microscopy

Cells were fixed in 2.5% Glutaralaldehyde in PBS over night at 4^0C, washed and post fixed in 1% OsO4 in PBS for 2h at 4^0C. Dehydration was carried out in graded ethanol followed by embedding in Glycid ether. Thin sections were mounted on Formvar/Carbon coated grids, stained with uranyl acetate and lead citrate and examined in Jeol 1400 Plus, transmission electron microscope (Jeol, Tokyo, Japan). Images were captured using SIS Megaview III and iTEM the Tem imaging platform (Olympus, Tokyo, Japan). All mice and human measurements and calculations in EM captures, were performed using Fiji open source platform for biological Image analysis. Analysis of patients bone marrow was approved by the IRB committee of Rabin Medical Center (approval 0840-18-RMC).

References

1. Ogilvy S, Metcalf D, Gibson L, Bath ML, Harris AW, Adams JM. Promoter elements of vav drive transgene expression in vivo throughout the hematopoietic compartment. *Blood*. 1999;94:1855-1863.
2. Thoms JA, Birger Y, Foster S, et al. ERG promotes T-acute lymphoblastic leukemia and is transcriptionally regulated in leukemic cells by a stem cell enhancer. *Blood*. 2011;117:7079-7089.
3. Fu YK, Tan Y, Wu B, et al. Gata2-L359V impairs primitive and definitive hematopoiesis and blocks cell differentiation in murine chronic myelogenous leukemia model. *Cell death & disease*. 2021;12:568.
4. Tsai FY, Keller G, Kuo FC, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*. 1994;371:221-226.